Evaluating the police service quality for handling traffic crash reporting

A combined MCDA and LCA approach

Janstrup, Kira Hyldekær; Kaplan, Sigal; Barfod, Michael Bruhn; Prato, Carlo Giacomo

Published in:
Policing: An International Journal

Link to article, DOI:
10.1108/PIJPSM-03-2016-0032

Publication date:
2017

Document Version
Peer reviewed version

Evaluating the police service quality for handling traffic crash reporting: A combined MCDA and LCA approach

Abstract

Purpose - The phenomenon of traffic crash under-reporting has been extensively documented in terms of its extent, but not equally analysed in terms of its reasons. As police distrust has been recently identified as a major reason for crash under-reporting, this study looks at the police service quality for handling the reporting of traffic crashes.

Design/methodology/approach – This study introduces a novel approach to evaluate service quality that combines Multi-Criteria Decision Analysis (MDCA) with Latent Class Analysis (LCA). Moreover, this study presents the design of a web-based survey on the basis of the SERVQUAL approach to detecting strengths, opportunities and threats with crash reporting to the police at a strategic level. Transportation stakeholders with an interest in traffic safety in Denmark participated in the survey that yielded 86 complete responses.

Findings - The novel approach was successfully applied and its implementation demonstrated the usefulness of the tool even in countries with a high police service. Results showed that the participating stakeholders perceived human factors as more important than physical factors in order to increase the crash reporting, with responsiveness as the most important and tangibles as the least important dimensions. Nevertheless, most stakeholders viewed a mixture of human and physical factors as crucial to increase crash reporting rates.

Originality/value – This study advances the knowledge about police service quality with a novel expert-based decision support tool based on SERVQUAL, MDCA and LCA, demonstrates its applicability in countries with a high police service, and opportunities and barriers for increasing the crash reporting rate.

Keywords Traffic crash reporting, police distrust, SERVQUAL, multi-criteria decision analysis, latent class analysis.

1. Introduction

Due to the reliance on traffic crash data from police reports for identifying human, vehicular and infrastructure factors underlying crash frequency and severity, concerns regarding their trustworthiness have been raised and questions about the under-reporting phenomenon have been asked (e.g., Elvik and Mysen, 1999; Farmer, 2003; Yannis et al., 2014). Studies comparing police and hospital crash databases have found high under-reporting rates in police data, with considerable selection bias related to road user type and crash severity (e.g., Elvik and Mysen, 1999; Yannis et al., 2014; Watson et al., 2015; Janstrup et al., 2016).

Perceived reporting usefulness and police distrust play a role in crash under-reporting (Amoros et al., 2007; Kaplan et al, 2016). Police distrust, namely a psychological state with negative feelings towards the police as an institution, is relevant to crash reporting as the police assumes the dual role of service provider and authority. Police distrust was mentioned as an impeding factor to incident reporting in cases involving domestic violence and sexual assault (Jones et al., 2009), was associated with low crime reporting rates in countries with high societal trust such as Scandinavian countries (Kääriäinen and Siren, 2011), and was related to perceived crash reporting usefulness among cyclists in Denmark (Kaplan et al., 2016).

Aware that achieving public safety and security goals depends on establishing trust in the community via their roles as public service provides and law enforcement agencies, police agencies have become more consumer-oriented and have increased their service quality in terms of accessibility and staff commitment (Drummond et al., 2000; Cukier et al., 2012). As service quality is considered essential for establishing trust in both the private and the public sector (e.g., Brysland and Curry, 2001; Kumar et al., 2009; Catulli, 2012), police agencies have used the SERVQUAL tool (Parasuraman et al., 1985) to identify the main components of police service quality (Mastrofski, 1999; Donnelly et al., 2006; Maguire and Johnson, 2010; Akhtar et al., 2011), with the main focus on either the perceived police service quality by the general public or the difference between expected and perceived service quality.

This study extends the body of knowledge on police service quality by focusing on traffic crash reporting, an important issue given that traffic crash reporting constitutes a large share of people's incident reporting to the police (Maguire and Johnson, 2010) and requires an improvement in quantity and quality. Specifically, this study offers a novel expert-based framework for evaluating police service quality in traffic crash reporting, detecting opportunities

and barriers to increasing crash reporting rates, and prioritizing service quality issues. The proposed tool combines the SERVQUAL model with an expert-based multi-criteria decision analysis (MCDA) and a latent class analysis (LCA) approach. The expert-based MCDA approach has been used previously for the evaluation of transport projects, policies and investments (e.g., Barfod, 2012; Wang et al., 2014; Macharis and Bernardini, 2015), but has been rarely used for service quality assessment with the only example in the healthcare sector (Oddershede et al., 2014). This study is novel in applying MCDA for police service quality assessment, and furthermore in considering LCA as a multi-dimensional approach to clustering stakeholders while considering their opinion and socio-economic heterogeneity. The expected impact of this study are insights on the perceived police service quality of handling crash reporting and highlights of the areas of improvement that would benefit the reporting rates the most.

The remainder of the paper introduces the three elements of the proposed tool (i.e. SERVQUAL, MCDA and LCA) and demonstrates its application to a country with high police service quality (Denmark). Then, the paper illustrates the results in terms of strengths, opportunities and barriers to improving crash reporting, and last presents major findings and concluding remarks.

2. Methodology

The proposed framework integrates SERVQUAL, MCDA and LCA for evaluating the police service quality in handling traffic crash reporting. The method is easily transferable to service quality evaluations in other fields, for example public transport, and its conceptual representation is depicted in figure 1.

Figure 1. Conceptual framework

2.1. SERVQUAL

At the core of the framework are Mastrosfki's six domains of performance (Mastrofski, 1999) derived from the SERVQUAL framework (Parasuraman et al., 1985). While the original framework evaluates service quality in the private sector across ten possible dimensions (reliability, responsiveness, competence, access, courtesy, communication, credibility, security, understanding, tangibles), Mastrofski (1999) identified six dimensions relevant to police agencies in their service provision to the general public: attentiveness, reliability, responsiveness,

competence, caring, and fairness. Mastrofski's dimensions have been used to evaluate the perceived police service quality in several countries, including U.K. (Donnelly et al., 2006), U.S. (Maguire and Johnson, 2010), Taiwan (Chu et al., 2010), and Pakistan (Akhtar et al., 2011).

As noted by Maguire and Johnson (2010) and Donnelly et al. (2006), existing studies have two major limitations: (i) the evaluations of the perceived service quality concerned a wide range of incidents and disregarded differences across incidents in terms of severity, occurrence probability, target population, required communication skills, and expertise; (ii) the evaluations focused on the perceptions of the general public or of the police, which opens to the possibility of strategic response bias, emotional bias and layman opinion. The proposed framework solves both these limitations: (i) this study focuses on the service quality needs for traffic crash reporting while considering the unique nature of traffic crashes that are unintentional violent incidents resulting in a range of consequences from damage only to fatal injuries, are relatively common among the general population, and in which the party at fault is often difficult to determine; (ii) this study takes an expert-based perspective to remove possible biases.

Given these premises, the framework postulates that the following dimensions impact police service quality when handling traffic crash reporting.

Accessibility is the perceived ease of contacting and reaching the police when reporting a traffic crash, measured in this study as short waiting for the police at the crash scene, short travel time to the police station, and attentive response to the public. Increasing the perception of the police as available and accessible boosts the perceived usefulness of crash reporting, as crash under-reporting is related to delays in the police arrival to the crash scene, concerns of disturbing the police, and doubts about the ability of the police to help (Kaplan et al., 2016).

Tangibles are the appearance and functionality of interior and exterior design of police stations, measured in this study as cleanliness of waiting environments, quietness of relaxed surroundings, and availability of computer-based documentation. Improving appearance and functionality of facilities, equipment and communication material (Donnelly et al., 2006) enhances the feeling of trust related to the perception of authority, approachability and professionalism (Clinton and Devlin, 2011).

Responsiveness is the willingness to help customers and provide prompt service (Donnelly et al., 2006), measured in this study as short waiting times, quick reporting procedures, clear information provision, effective handling of the public, and allocation of sufficient time for

reporting. Increased responsiveness augments crash reporting rates, as the perception of time-consuming, non-flexible and cumbersome reporting systems are major barriers to incident reporting (e.g., Kongsvik et al., 2012; Lindsay et al., 2012; Winswold-Prang and Jelsness-Jørgensen, 2014; Kaplan et al., 2016).

Caring is the individualised attention an organisation provides to its customers (Donnelly et al., 2006), measured in this study as general friendliness, positive attitude, and supportive non-violent communication focused on understanding the needs of people who arrive to report traffic crashes. Friendliness and readiness to help increase the trust in the police (Kaplan et al., 2016) and help coping with the self-blame and guilt feelings often experienced by traffic crash victim (Peltzer and Renner, 2004) and the concerns about family distress and social image that are barriers for traffic crash reporting intentions (Kaplan et al., 2016).

Fairness is being fair with respect to the procedure leading to an outcome and the interaction between decision-makers and individuals affected by the decision (Maguire and Johnson, 2006), measured in this study as allocating sufficient time to the crash report, being precise in the registration, and being neutral and fair with the crash victims. Fairness is related to trust (Maguire and Johnson, 2006) and increases crash reporting intentions, as these are related to preferences to allocate time to activities other than the civic duty of crash reporting (Kaplan et al., 2016).

2.2. MCDA

MCDA considers explicitly multiple criteria and applies the decomposition approach for identifying strengths, opportunities, threats and possible policies in police service in order to improve traffic crash reporting. The proposed framework combines the SERVQUAL scale with the Simple Multi-Attribute Rating Technique (SMART, Von Winterfeldt and Edwards, 1986) that evaluates a finite number of decision alternatives with respect to a finite number of performance criteria. Specifically, SMART relies on the additive value function model and assigns direct scores to alternatives and direct weights to criteria using a scale tailored to the problem at hand (Barfod and Salling, 2015).

The input to the MCDA is obtained from an expert panel (e.g., stakeholders, decision-makers) and can be elicited with standard procedures (see, e.g., Shiftan et al., 2003; Wang et al., 2014). This study uses the Delphi-method (Sackman, 1974) to stakeholders in the transport field (e.g., researchers, authorities, consultants, NGO representatives, suppliers) who participated in a

pilot test and the final survey. The choice of the Delphi-method is in line with numerous studies concerning scientific and technological forecasting, although examples exist also in social science with non-expert panels.

After selecting the expert panel, the framework consisted of four consecutive steps: (i) preparing a survey that elicits items for each of the five dimensions from the SERVQUAL scale (i.e., accessibility, tangibles, responsiveness, caring and fairness) to detect strengths, opportunities and threats dimensions associated with handling traffic crash reporting; (ii) applying the SMART technique to rate challenges and opportunities of each dimension while ranking the dimensions in a subjective order of preference and rating the overall impact; (iii) applying the SMART technique to elicit criteria weights for each item of each dimension; (iv) asking the experts to reflect about the improvement possibilities for crash reporting if the issues identified in the five dimensions are solved. It should be noted that the second and third step take the decomposition approach by evaluating the dimensions as separate items according to the SMART, and that the large-scale survey suggested to use a 5-point Likert-scale for the third step. It should be noted that SMART relies on the linear additive model and hence the overall value over a dimension is equal to the sum of the performance scores multiplied by the weights. Last, it should be also noted that the last step gives the respondents the opportunity of reflecting about their answers in the previous parts of the survey.

2.3. LCA

The result of an expert-based evaluation may be consensus or contrasting opinions, and a single solution is not mandatory (Shiftan et al., 2003). Having a range of solutions is actually important as a sensitivity measure that increases the MCDA robustness by mapping sources of uncertainty (Barfod and Salling, 2015). This study proposes a multi-dimensional model-based technique to differentiate stakeholders according to their opinions and their social and professional characteristics with the aim of identifying a range of alternative solutions.

LCA (Magidson and Vermunt, 2002) is an unsupervised learning approach without a priori hypothesis about the clusters' composition according to the stakeholders' social and professional characteristics. LCA was conceived over four decades ago, but only recently renewed interest in its application and advances in computational capabilities led to wide-spread applications in a diversity of social science studies (Lanza et al., 2007). LCA is preferred over traditional clustering approaches because of its ability to account for overlap across clusters, its reliance on

an underlying statistical model that allows calculating cluster probabilities for new cases, and its goodness-of-fit criteria that guide the decision about the number of clusters (Kaplan and Prato, 2013; Weiss et al., 2015). LCA technique was performed in the current study by using the SAS procedure developed by Lanza et al. (2007).

In this study, LCA classifies stakeholders into *C* latent classes of unknown size and composition. Given a vector of stakeholders that gave answers to *M* items ($Y_i = y_1,\ldots,y_M$), and given the vector ($Y_i=Y_{i1},\ldots,Y_{iM}$) of values for stakeholder *i* over each item, the probability of stakeholder *i* to be in class *k* is expressed as (Lanza et al., 2007):

$$p\left(Y_i \mid \theta\right) = \sum_{k=1}^{K} P\left(C_k\right) p\left(Y_i \mid C_k, \theta_k\right) \tag{1}$$

where *K* is the number of classes, $P(C_k)$ is the prevalence of class C_k in the dataset, $p(Y_i \mid C_k,\theta_k)$ is the conditional multivariate probability that a stakeholder in class C_k would be characterized by Y_i, and θ_k is a vector of unknown parameters to be estimated. In order to derive an estimable model formulation with reasonable parametric complexity, this study makes the following assumptions: (i) every item in the model is defined as an ordinal indicator with R_m possible responses; (ii) the stakeholders' answers are assumed to be independent; (iii) the categorical indicators are assumed to be independent within a latent class, namely the within-class covariance matrix is assumed to be diagonal; (iv) the categorical indicators are assumed to be endogenous indicators of the latent class and no covariates are employed to predict class membership. Under these assumptions, the LCA model is formulated as (Lanza et al., 2007):

$$p\left(Y_i \mid \theta\right) = \sum_{k=1}^{K} \pi_k \prod_{m=1}^{M} \prod_{r_m=1}^{R_m} \theta_{mr_m|k}^{I(y_{im}=r_m)} \tag{2}$$

where *I* is an indicator function that equals 1 if y_{im} equals r_m and *0* otherwise, π_k are the class membership probabilities to be estimated, and θ_{mrm} are indicator response probabilities conditional on the latent class membership to be estimated.

Maximum-likelihood using the expectation-maximization (EM) algorithm and the maximum absolute deviation (MAD) convergence criterion is used to estimate the LCA parameters. As the likelihood function is not necessarily concave, EM was repeated with 50 sets of starting values to guarantee the convergence to the global maximum rather than local maxima. The number of clusters was chosen according to the Bayesian Information Criterion (BIC) as well as the ability to distinguish and interpret clusters of non-negligible size (Lanza et al., 2007).

3. Data

3.1. Survey design

Data were collected by designing a web-based survey upon the SERVQUAL scale developed in this study to elicit stakeholders' perceptions regarding the police service quality for handling traffic crash reporting.

Five parts composed the survey and elicited: (i) stakeholders' socio-economic characteristics (e.g., age, gender and workplace location) and professional experience with traffic safety and crash data analysis; (ii) stakeholders' personal experience as a private person visiting the police for crash reporting or other reasons in the last five years; (iii) stakeholders' perceptions regarding the importance of reporting crashes to the police and the importance of the five dimensions for handling traffic crash reporting; (iv) stakeholders' confidence level in answering questions about the police and rating of the importance of each question within the five dimensions for handling traffic crash reporting; (v) stakeholders' belief in the ability to improve crash reporting rates as a result of police service quality improvements. The questions were rated on a 5-point Likert scale.

3.2. Survey administration

A preliminary version of the survey was administered to the expert panel from the transport field at the Danish Transport Conference (Trafikdage) at Aalborg University. The low completion rate (39%) suggested to change and shorten the survey, with the two main improvements in the structure rather than the content: (i) a reduction in each item from 11 to 5 points in the Likert-scale; (ii) a 50% reduction in the number of questions by focusing on service quality improvements instead of current situation, service quality and ease of improving. The final survey was administered to researchers (e.g., DTU Transport, Aalborg University), authorities (e.g., Danish Road Directorate, Danish Police, municipalities), consultants (e.g., Cowi, Trafitec), NGO's (e.g. Danish Council for Safer Traffic, Danish Cyclist's Federation), and transport suppliers (e.g., Danish Interurban Railways, Danish Railways).

3.3. Sample characteristics

While the preliminary survey was completed by 39% of the sample, mainly male stakeholders working for transport authorities, the final survey was completed by 72% of 125 participants. The sample was composed of mainly males (59%) and workers from transport authorities (41%), constituting the largest interest group. Among the participants, 27% was between 30 and 39 years old, and 28% were between 50 and 59% years old. The vast majority of the respondents

(94%) declared an interest in traffic safety and almost a third (30%) stated to have visited a police station at least once in the last five years. High and very high confidence in the ability to evaluate police service quality for traffic crash reporting was indicated by 29% of the participants, and medium confidence was declared by another 36%.

4. Results

4.1. Perceptions regarding current service quality

The preliminary survey provided a sense of the evaluation of the current service quality and the perceived ease of improvement for handling traffic crash reporting, but length reasons motivated its removal in the final survey. The scale for the service quality ranged between 0 (non-existent) and 10 (excellent), while the scale for perceived ease of improving ranged from 0 (impossible) to 10 (already perfect).

Table 1 summarizes the answers to these questions. Results suggested that the stakeholders perceived a good service quality at police stations and a general satisfaction with their handling of traffic crash perception. The differences in the perceived service quality across the five dimensions were quite small, indicating the lack of evident lacunas.

High score on the ***caring*** dimension was not surprising in a Danish context with high societal trust, and suggested the opportunity for police agencies in other countries to explore the transferability of the Danish communication model in handling traffic crash reporting. ***Caring*** was the dimension with the highest level of service and moderate ease of improvement, while ***tangibles*** was the dimension with the lowest service quality and the highest ease of improvement. Specific items that scored low in terms of service quality were travel time to police stations (***accessibility***), computer resources (***tangibles)***, waiting time (***responsiveness***), and time allocated for crash reporting (***fairness***), with computer resources rated as easy to resolve and travel time to police stations as difficult to improve.

Table I. Average value and standard deviation for the survey questions in the preliminary survey

4.2. Importance of service quality dimensions for improving crash reporting rates

The final survey provided the input to SMART to obtain an understanding of the responses and a weighing of the importance of each dimension. The scores represent the importance for each item in the five dimensions when reporting a traffic crash to the police and they are weighted according to the importance of each dimension. Table 2 presents the weighted average score for

each question in the five dimensions across the whole sample, with the resulting scores normalized to sum to 1.

Table II. Survey questions and the five dimensions

As items in each dimension had similar scores, the evaluation of service quality can be dimension-based rather than item-based. ***Responsiveness***, ***caring*** and ***fairness*** were perceived as more important than ***accessibility*** and ***tangibles***, indicating that human factors of communication and interaction were viewed as more important than physical factors to increase crash reporting rates. When looking at the item detail, travel time to police stations and their ambiance were the least important, while sufficient time for traffic crash reporting and solving problems arisen in the process were the most important.

4.3. Heterogeneity of stakeholders' perspectives

LCA identified clusters for the responses received from the expert panel by using the categorical indicators of their characteristics and the weighted importance evaluations of their answers to each item of each dimension in the SERVQUAL scale. The BIC values for 7 clusters and the entropy of the solution equal to 1.00 that indicates a very high certainty in the classification indicated the existence of 7 groups of stakeholders that were sorted according to differences and similarities in their prevalent features (i.e., at least 55% of the observations share the same feature in the cluster).

Table 3 describes for each cluster the stakeholders' characteristics, their professional and personal experience, and their perceptions of the importance of each item.

Table III. Latent class characteristics (percentages of cluster observation)

Cluster C1 includes mainly professionals over 40 years old who work as researchers in Zealand and focus on office work in the safety area. They have experience with writing and analysing safety reports, they have no experience on the field, they perceive contacting the police for traffic crashes as important, and they self-assess having low certainty in answering the survey. They rank all ***fairness*** items as of low importance, part of ***tangibles*** and ***accessibility*** items as of medium-to-high importance, and all ***responsiveness*** and ***caring*** items as of high importance.

Cluster C2 comprises male professionals working in public authorities or consultancy companies in Zealand or Southern Denmark. They have experience with both office and field

work in traffic safety as they have worked extensively with reading and analysing safety reports as well as inspections, revisions and safety plans. Most professionals in this clusters have experience with crash data and black spot analysis, they perceive contacting the police in traffic crashes as important, and they are highly certain of their answers in the survey. They rank all ***caring*** items as of low importance, part of ***tangibles*** and ***accessibility*** items as of medium-to-high importance, and all ***responsiveness*** and ***fairness*** items as of high importance.

Cluster C3 contains professionals who are 30-49 years old and work mainly in Zealand. They have an interest in traffic safety, although they have mostly limited experience besides reading reports, they perceive that contacting the police in the case of a traffic crash is important, and they have low certainty about their answers to the survey. They rank all ***tangibles*** items as of low-to-medium importance, and all ***accessibility***, ***caring***, ***responsiveness*** and ***fairness*** items as of high importance.

Cluster C4 is made of professionals who are 40-59 years old and work mainly as consultants in Zealand. They have experience with both office and field work in traffic safety, they have mostly experience with crash data and black spot analysis, they perceive as important to contact the police for traffic crash reporting, and they have medium certainty in their answers. They rank all ***tangibles*** items as of low importance, all ***fairness*** items as of medium importance, all ***caring*** and most ***accessibility*** items as of medium-to-high importance, and all ***responsiveness*** items as of high importance.

Cluster C5 includes two respondents with a diametral difference of opinions, and hence the results are not discussed.

Cluster C6 comprises female professionals who work in public authorities and have mainly experience in reading safety reports. They perceive that contacting the police when a traffic crash occurs is important and they self-assess their certainty in answering the survey as medium. They rank all the items across the five dimensions as of medium importance.

Cluster C7 contains professionals over 50 years old, mainly males working in either public authorities or consultancy companies. They disagree on the importance of calling the police when a traffic crash happens, they have experience in traffic safety planning and reading reports, and they evaluate their certainty in answering the survey as medium-to-high. They rank all ***accessibility*** and part of the ***tangibles*** items as of low-to-medium importance, all ***responsiveness***

and ***caring*** items as of medium importance, all ***caring*** and most ***accessibility*** items as of medium-to-high importance, and most ***fairness*** items as of high importance.

4.4. Possibility for improving traffic crash reporting

Table 4 presents the stakeholders' perceptions of the possibility of improving police service quality for handling traffic crash reporting as an overall estimate for each cluster as well as a total estimate for the sample. Table 5 lists the perceived increase in the crash reporting rate to the police, provided the necessary improvements along each quality of service dimension.

Table IV. The possibility for improving traffic crash reporting (percentages of the cluster observations)

Table V. The perceived increase in crash reporting rate to the police, provided service quality improvement along the specified dimensions (percentages of the cluster observations)

At the sample level, 30.2% of the experts perceive that improving the crash reporting rates is highly possible, while only 9.3% believe that it has low possibility. At the cluster level, cluster C2 is the most optimistic, followed by cluster C6, while cluster C7 is the most pessimistic about possible improvements to the reporting rates.

Looking at the entire sample across the five dimensions, 20-30% of the experts perceive that a particular dimension would be associated with an increase in crash reporting rates. The most promising dimension is ***responsiveness*** (35% of the experts assign high increase), while the least promising dimension is ***accessibility*** (37% of the experts assign low increase).

Looking at differences across clusters, ***accessibility*** is associated with a low increase in crash reporting by experts in clusters C2 and C7, and with a high increase by experts in clusters C4 and C6. With respect to the total sample, a higher percent of experts in cluster C6 considers ***tangibles*** as associated with a high increase in crash reporting rates, a higher share of experts in clusters C3 and C6 deems ***caring*** as related to the same high increase, and a higher percent of experts in cluster C2 and C3 judges ***responsiveness*** as linked to the same high increase. Also with respect to the total sample, a higher percent of experts in cluster C2 and C3 evaluates ***fairness*** as associated to a medium-to-high increase in the reporting rate, and a higher share of experts in cluster C7 considers ***fairness*** related to a low increase in the reporting rate.

5. Conclusions

This study proposed a framework for evaluating police service quality when handling traffic crash reporting. Framed around the Mastrofski's dimensions from the SERVQUAL scale, the

integration of MCDA and LCA allowed investigating opportunities and barriers for handling traffic crash reporting and highlighting avenues for improvements to the reporting rates. Findings from this study extend existing findings about police service quality (Mastrofski, 1999; Donnelly et al., 2006; Chu et al., 2010; Maguire and Johnson, 2010; Akhtar et al., 2011) and open options for service quality evaluation in other fields.

The successful implementation of the proposed framework showed that there is a general satisfaction with the service quality in the current situation, with caring scoring the highest in a Danish society with a high societal trust in the police. This is an important result as it opens the investigation of the transferability of the Danish communication model in handling traffic crash reporting to other countries. Also, findings from the implementation of the framework showed that stakeholders view human factors (i.e., communication, interaction) as more important than physical factors in order to increase crash reporting rates. Namely, stakeholders evaluated ***responsiveness***, ***caring*** and ***fairness*** as more important than ***accessibility*** and ***tangibles.*** Noticeably, although human factors scored quite high for the current situation, the results highlight the need for further improvements.

The implementation of clustering analysis according to the stakeholders' social and professional characteristics revealed differences of opinions about where and how much to invest in the different dimensions. For example, the clusters of experts perceiving the possibility for a high improvement of reporting rates within ***accessibility*** and ***tangibles*** mainly consists of females working as authorities. However, there are also some agreements across clusters: (i) the experts in all clusters agree that ***responsiveness*** is the most important dimension to be improved and that this improvement will yield the highest increase in crash reporting rates, and (ii) only one cluster (C7) considers only human factors as relevant, while all others see the improvements coming from a combination of measures relating to human and physical factors.

Overall, the advantages of the proposed method for evaluating service quality consist in the introduction of a weighting system for the improvement rates, the proposition of a reflection procedure about the confidence in the answers, and the ability to represent overlap across clusters. A practical advantage is that the method proves useful even in a country with a high police service quality and opens avenues for possible new results for service quality evaluations. A further research direction could be the adaptation and application of the proposed method also

to other service fields combining physical facilities and human interaction such as the aviation industry, tourism, and public transport provision.

References

Akhtar, A., Rafiq, S., Asif, A., Saeed, A., and Kashif, M. (2011), "Public perceptions of police service quality: empirical evidence from Pakistan", *International Journal of Police Science & Management*, 14, 97-106.

Amoros, E., Martin, J., and Laumon, B. (2007), "Estimating non-fatal road casualties in a large French County, using the capture-recapture method", *Accident Analysis and Prevention*, 39, 483-490.

Barfod, M.B. (2012), "An MCDA approach for the selection of bike projects based on structuring and appraising activities", *European Journal of Operational Research*, 218, 810-818.

Barfod, M.B., and Salling, K.B. (2015), "New composite decision support framework for strategic and sustainable transport appraisals", *Transportation Research Part A*, 72, 1-15.

Brysland, A., and Curry, A. (2001), "Service improvements in public services using SERVQUAL", *Managing Service Quality: An International Journal*, 11, 389-401.

Catulli, M. (2012), "What uncertainty? Further insight into why consumers might be distrustful of product service systems", *Journal of Manufacturing Technology Management*, 23, 780-793.

Chu, K., Li, C., and Chiu, P. (2010), "A study of the quality of police services in exercising public power in Taiwan", *Total Quality Management & Business Excellence*, 21, 537-551.

Clinton, A., and Devlin, A.S. (2011), "Is this really a police station? Police department exteriors and judgments of authority, professionalism, and approachability", *Journal of Environmental Psychology*, 31, 393-406.

Cukier, W., Barkel, E., Vaughan, T., and Gekas, G. (2012), "Quality assurance in Canadian police services", *The TQM Journal*, 24, 295-309.

Donnelly, M., Kerr, N.J., Rimmer, R., and Shiu, E.M. (2006), "Assessing the quality of police services using SERVQUAL", *Policing: An International Journal of Police Strategies & Management*, 29, 92-105.

Drummond, G., Ensor, J., Laing, A., and Richardson, N. (2000), "Market orientation applied to police service strategies", *International Journal of Public Sector Management*, 13, 571-587.

Elvik, R., and Mysen, A.B. (1999), "Incomplete crash reporting: meta-analysis of studies made in 13 countries", *Transportation Research Record*, 1665, 33-140.

Farmer, C.M. (2003), "Responsiveness of police-reported information for determining crash and injury severity", *Traffic Injury Prevention*, 4, 38-44.

Janstrup, K.H., Kaplan, S., Hels, T., Lauritsen, J., and Prato, C.G. (2016), "Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics", *Traffic Injury Prevention*, doi:10.1080/15389588.2015.1128533.

Jones, J.S., Alexander, C., Wynn, B.N., Rossman, L., and Dunnuck, C. (2009), "Why women don't report sexual assault to the police: the influence of psychosocial variables and traumatic injury", *The Journal of Emergency Medicine*, 36, 417-424.

Kaplan, S., and Prato C.G. (2013), "Cyclist-motorist crash patterns in Denmark: A latent class clustering approach", *Traffic Injury Prevention*, 14, 725-733.

Kaplan, S., Janstrup, K.H., and Prato, C.G. (2016), "To report or not to report: Applying the Theory of Planned Behavior to understand the intentions to report cycling incidents by young adults", *Proceedings of the 95th Annual Meeting of the Transportation Research Board*, Washington, D.C.

Kääriäinen, J., and Sirén, R. (2011), "Trust in the police, generalized trust and reporting crime", *European Journal of Criminology*, 8, 65-81.

Kongsvik, T., Fenstad, J., and Wendelborg, C. (2012), "Between a rock and a hard place: Accident and near-miss reporting on offshore service vessels", *Safety Science*, 50, 1839-1846.

Kumar, M., Kee, F.T., and Manshor, A.T. (2009), "Determining the relative importance of critical factors in delivering service quality of banks", *Managing Service Quality: An International Journal*, 19, 211-228.

Lanza, S.T., Collins, L.M., Lemmon, D.R., and Schafer, J.L. (2007), "PROC LCA: A SAS procedure for latent class analysis", *Structural Equation Modeling*, 14, 671-694.

Lindsay, P., Sandall, J., and Humphrey, C. (2012), "The social dimensions of safety incident reporting in maternity care: The influence of working relationships and group processes", *Social Science & Medicine*, 75, 1793-1799.

Macharis, C., and Bernardini, A. (2015), “Reviewing the use of Multi-Criteria Decision Analysis for the evaluation of transport projects: Time for a multi-actor approach”, *Transport Policy*, 37, 177-186.

Magidson, J., and Vermunt, J.K. (2002), “Latent class models for clustering: A comparison with K-means”, *Canadian Journal of Marketing Research*, 20, 37-44.

Maguire, E.R., and Johnson, D. (2010), “Measuring public perceptions of the police”, *Policing: An International Journal of Police Strategies & Management*, 33, 703-730.

Mastrofski, S. (1999), “Policing for people”, Ideas in American Policing lecture at the Police Foundation, Washington, D.C. Available at: http://www.policefoundation.org/wp-content/uploads/2015/06/Mastrofski-1999-Policing-For-People.pdf

Oddershede, A.M., Cordova, F.M., Carrasco, R.A., and Watkins, F.J. (2014), “Decision model for assessing healthcare ICT support implications: user perception”, *International Journal of Computers Communications & Control*, 9, 593-601.

Parasuraman, A., Zeithaml, V.A., and Berry, L.L. (1985), “A conceptual model of service quality and its implications for future research”, *Journal of Marketing*, 49, 41-50.

Peltzer, K., and Renner, W. (2004), “Psychosocial correlates of the impact of road traffic accidents among South African drivers and passengers”, *Accident Analysis and Prevention*, 36, 367-374.

Sackman, H. (1974), “Delphi assessment: expert opinion, Forecasting and group process”, RAND report R-1283-PR, 7-8.

Shiftan, Y., Kaplan, S., and Hakkert S. (2003), “Scenario building as a tool for planning a sustainable transportation system”, *Transportation Research Part D*, 8, 323-342.

Von Winterfeldt, D., and Edwards, W. (1986), “*Decision Analysis and Behavioural Research*”, Cambridge University Press, New York, U.S.

Wang, Y., Monzon, A., Di Ciommo, F., and Kaplan, S. (2014), “An integrated transport planning framework involving a combined utility-regret approach”, *Transportation Research Record*, 2429, 59-66.

Watson, A., Watson, B., and Vallmuur, K. (2015), “Estimating under-reporting of road crash injuries to police using multiple linked data collections”, *Accident Analysis and Prevention*, 83, 18-25.

Weiss, H.B., Kaplan, S., and Prato, C.G. (2015), “Fatal and serious road crashes involving young New Zealand drivers: a latent class clustering approach”, *International Journal of Injury Control and Safety Promotion*, doi: 10.1080/17457300.2015.1056807.

Winsvold-Prang, I., and Jelsness-Jørgensen, L.P. (2014), “Should I report? A qualitative study of barriers to incident reporting among nurses working in nursing homes”, *Geriatric Nursing*, 35, 441-447.

Yannis, G., Papadimitriou, E., Chaziris, A., and Broughton, J. (2014), “Modeling road crash injury under-reporting in Europe”, *European Transport Research Review*, 6, 425-438.